Resources Policy 72 (2021) 102117

Contents lists available at ScienceDirect

Resources Policy

journal homepage: http://www.elsevier.com/locate/resourpol

Impact of Corporate Social Responsibility in mining industries

Adrià Pons [b], Carla Vintrò [b], Josep Rius [b,*], Jordi Vilaplana [a]

[a] Dept. of Computer Science, University of Lleida, Av. Jaume II 69, 25001, LLeida, Spain
[b] Dept. of Business Administration, University of Lleida, Av. Jaume II 71, 25001, LLeida, Spain

A R T I C L E I N F O

Keywords:
CSR
Mining
Twitter
Big data
Sentiment analysis

A B S T R A C T

The mining sector plays a fundamental role in the global economy since it provides vital raw materials and energy for a large number of industries but mining activities are commonly criticized due to the effects on workers' health and local communities, and are seen as a threat to society in general. For these reasons, companies have to deal with the compatibility between the productive activity and the Corporate Social Responsibility (CSR). Under this context, this paper aims to examine CSR communication in the mining sector on Twitter and identify the main topics of CSR and the main participants in the creation of content. A software framework has been developed in order to manage big text data in the study of CSR and sentiment computational content analysis to classify the tweets as positive, negative or neutral. Twitter was chosen since it is one of the three most commonly used social network sites by mining companies for information sharing and stakeholder engagement. The data collection period was from February 2019 to December 2019 and it resulted in 2,000,000 Twitter posts. The results show the CSR debate is increasingly growing in developing countries and in countries with a bad reputation of environmental and health mining conditions. The debate is mainly centered on the impact of mining industries on the land, and many stakeholders advocate for more environmental responsibility since they consider that mining activities are still controversial.

Also, results reveal that the mining sector should improve the CSR disclosure and adopt a stakeholder engagement strategy grounded in the corporate stakeholder relationship perspective, the two-way symmetrical communication, and the dialogic theory of public relations. Social network sites such as Twitter can lead to positive outcomes for companies since they are not only a way to advertise what the company does with regards to CSR, but also to receive input from other for their CSR activities.

1. Introduction

Over the past several years, stakeholder relationship management has played an important role in strategic leadership and within the scope of business models. Companies that are seeking for sustainability and ethical management of their activities need to know and grasp their interest groups, enhance dialogue, meet their demands and expectations as well as be transparent when reporting back their actions to generate real value for their different stakeholders. Hence, it is necessary to measure the impact of the actions undertaken by the com-panies together with interpretations and feelings being generated towards each interest group, and thus be able to redefine priorities and responsibilities within the organization from a triple perspective: economic, social and environmental. Having greater insight into the real purpose of the organization, in other words justify the value and impact of the activity on society, introduce more sustainable practices to tackle current challenges and stakeholders' interests, identify essential products and services that define the core business and create true social value. This would yield useful information about which are the strategic priorities that companies should face from the Corporate Social Responsibility (hereinafter CSR) point of view. In this regard, environmental impact of economic activities as well as employee security and health conditions are two of the main worldwide concerns and have become an issue of great importance due to are part of the United Nation 2030 agenda for Sustainable Development Goals. These concerns have particular relevance in some sectors depending on the nature of its activities, such as the mining industry.

This sector plays a fundamental role in the global economy since it provides vital raw materials and energy for a large number of industries. Its activities have important economic, environmental, labor and social repercussions on local and global scales Escanciano et al. (2010). As stated by Mendes and Rodrigues, (2018) they have specific

* Corresponding author.
E-mail address: josep.riustorrento@udl.cat (J. Rius).

https://doi.org/10.1016/j.resourpol.2021.102117
Received 21 April 2020; Received in revised form 6 April 2021; Accepted 14 April 2021
Available online 29 April 2021
0301-4207/

characteristics because of their transitory nature, and are still commonly considered as a threat to the natural surroundings, with environmental effects on the air, water and soils. Moreover, mining activities are commonly criticized due to the effects on workers' health Sanmiquel et al. (2015) and local communities Raufflet et al. (2014), and are seen as a threat to society in general Mendes and Rodrigues (2018). Further, companies have to deal with the compatibility between the productive activity and the environmental Claver Cortés et al. (2004) and social protection Wheeler et al. (2002), which are pillars of sustainable development.

Research on CSR facilitates a strategy to face all these challenges and helps companies to undertake different actions in order to accept the responsibilities resulting from the impact of its activities on society and environment Vintro et al. (2012) under four dimensions; economic, legal, ethical and philanthropic Carroll (1991). CSR aims to connect the objective key results of the company with society and stakeholders' expectations.

The concept of CSR has gained importance within the academic community and has been an increasing studied subject in recent years, with particular significance in mining activities. As noted by Warhurst Warhurst (2001a), the main environmental disasters and human rights incidents that have increased public concern about CSR over the last 40 years have mainly taken place in the mining and petroleum industries Vintro et al. (2012). The discussion about CSR in the mining industry has emphasized its commitment to social responsibility as an emerging topic Mendes and Rodrigues (2018), and literature seems to indicate that mining companies have increased their environmental and social consciousness Zhu et al. (2010). In this sense, the first decade of the 21st century in particular has seen a renewed debate about mining and its sustainability.

M. Mudd (2010) owing to public concern about the current degradation of the environment Hilson and Murck (2000) and about social actions in the communities affected Hamann (2004). In the literature, there are several studies regarding the extensive interest on the topic of CSR which explore strategies of mining companies on sustainable issues; Sinding (2009); Mclaren et al. (1999); Warhurst and Noronha (2009); Hilson and Murck (2000); Warhurst (2001a); Hilson and Nayee (2002); Newbold (2006); Suppen et al. (2006); Van Zyl et al. (2007); Rodrigues da Silva Enríquez and Drummond (2007); M. Mudd (2010); Fonseca (2010); Dutta et al. (2012); Hassan and Ibrahim (2012); Vintro et al. (2014).

Thus, CSR is increasingly studied by different researchers and is a core strategy of many companies, but as noted by Chae and Park (2018) there are many needs and opportunities for future CSR research, one of those is the application of content analysis to extract themes, categories and sentiments from text corpora. Most of the published studies about content analysis in CSR (see for example Dahlsrud (2008) and Lee and Carroll (2011) have examined small samples of text data from traditional media including journal articles, CSR reports and newspapers, but very few (see for example (Chae and Park, 2018) have examined big amounts of data obtained from social media even though the analysis of big data helps researchers identify some latent knowledge or information to be incorporated in future decision-making Crane and Self (2014). Moreover, these previous studies have mainly used human coding or simple software-based methods in their content analysis (Chae and Park, 2018). When considering the research on CSR in the mining sector, the use of content analysis is even scarcer.

The arrival of big data has shown an improvement in the identification and analysis of perceptions and sentiments from stakeholders which are relevant for all sectors, but even more for those seen as potentially dangerous such as mining. Now, the discussions have mainly shifted to social media where either internal and external stakeholders express most of their comments and criticisms in relation to actions taken by companies and its consequences. There is large interest in studying sentiment analysis from these Stakeholder's comments and posts on social media using big data techniques in order to help companies redefine strategic priorities together with actions to be taken under a social responsibility and environmental framework, which goes beyond the question of economic results.

Under this context, this paper aims to examine CSR communication in the mining sector on Twitter and identify the main topics of CSR and the main participants in the creation of content. In other words, the results obtained will contribute to the development of CSR strategies for the mining industry and similarly other studies may be conducted in different sectors to redefine the CSR role and priorities. The study is sustained in how organizational research focused on social media may translate stakeholder engagement into organizational goals and create the basis of effective strategy development.

The research conducted uses an Application Programming Interface (API) to manage big text data in the study of CSR and sentiment computational content analysis to classify the tweets as positive, negative or neutral. To this aim, the authors developed a tool in Phyton, through the Application Programming Interface (API) of Twitter, the Phyton language, multiple scripts and MongoDB. Twitter was chosen since it is one of the three most commonly used social network sites by mining companies for information sharing and stakeholder engagement. The data collection period was from February 2019 to December 2019 and it resulted in 2,000,000 Twitter posts.

In the next section, the paper explores the current discussions on CSR communication and Social Media. The project architecture of the framework, describing its functionality and the algorithms implemented, is then introduced. Next, the results obtained are presented. Discussions and conclusions follow.

2. Theoretical framework: CSR communication and social media

CSR emphasizes the role of corporate communication in company-stakeholders engagement, according to the principles of transparency and open dialogue with diverse stakeholders to foster ethical and sustainable behavior Lim and Green-wood (2017). In recent years, CSR communication has become a keystone when constructing and sustaining the reputation of a company in the eyes of its stakeholders Türkel and Akan (2015), and it has drawn considerable attention form public relations and CSR researchers Lim and Greenwood (2017). Mining activities might be viewed as more socially and environmentally responsible if CSR practices were adopted. Vintro et al. (2012) concluded that mining companies understand the negative impact of their activities on the environment and they want to collaborate with both internal and external stakeholders.

Until recently, companies mainly communicated through media relations, CSR reports, or websites among other means. Currently, companies also use a wide range of electronic web 2.0 based applications that include social networks, blogs, and photo and video sharing platforms. That is, CSR communication has evolved from one-way communication to two-way communication, and social media has opened new opportunities for CSR related information dissemination and stakeholder relationship management Lattemann and Stieglitz (2007).

Social media can be understood as “Internet-based applications that carry consumer-generated content which encompasses media impressions created by consumers, typically informed by relevant experience, and archived or shared online for easy access by other impressionable consumers” Xiang and Gretzel (2010). Therefore, social media refers to online resources that people use to share content. It is a public channel for firms to engage stakeholders since it allows for rapid dissemination and exchange of information, and by this way enables enterprises to cultivate relationships with key stakeholders Kelleher (2007) and engage in conversations. Stakeholders are no longer simple receivers of information and can now engage in the creation and evaluation of content Dellarocas (2003).

Social media cover a wide range of Social Network Sites, that is, sites driven by user-participation and user-generated content Tredinnick (2006). From those, Facebook, Instagram and Twitter are three of the most commonly used by companies. Specifically, Twitter was launched in October 2006, and it has become the largest micro-blogging site on the Internet. Around 316 million daily active users around the globe post 500 million tweets per day Kühl et al. (2019). It allows users to broadcast real-time messages of 280 characters or less. Users may select other users to “follow”, repost (“retweet”) updates from that users, pose comments (@reply), and mark “likes”. It also functions as a social networking site in that users can connect and share information and is used in official public relations, advertising, and marketing campaigns Stelzner (2012). According to the literature, sentiment classification provides organizations with a tool to transform data into ‘actionable knowledge’ that decision maker can use in pursuit of improved organizational performance. Especially in case of machine learning based, the classification performance can directly depend on the quality of features obtained from a training dataset Cortes et al. (1995) Mitchell (1999) Kira and Rendell (1992). Most of existing sentiment classification studies provide simple performance comparison in terms of the “accuracy” of algorithms without considering the role of “data properties and setting” that may cause differences in the performance, however, recent and advanced theory proposed by Choi and Lee (2017) offers a new approach investigating the impact of linguistic properties of data such as word-count and size of training dataset on the performance of diverse algorithms which leads to the conclusion that the quality and properties of training dataset are of considerable importance for the performance of machine learning and classification accuracy, thus affecting the final results credibility.

Organizational research specific to Twitter has gained importance within the academic community during last years. Different authors have discussed what best practices of communication on Twitter may be boyd et al. (2010), Rybalko and Seltzer (2010), Marwick and Boyd (2011). For instance, Waters and Jamal Waters and Jamal (2011) have examined how non-profit organizations from the Philanthropy 200 communicate on Twitter. Xiang and Gretzel (2010) have analyzed how Twitter contributes to the development of the theory and practice of public relations. Rybalko and Seltzer (2010) have examined how Fortune 500 companies engage stakeholders using Twitter. And similarly, Barnes and E.M (2010) has studied how Fortune 500 companies use Twitter and found that only 35% of those use it, and only 24% of that are actively involved.

Different studies have analyzed organizational engagement through Twitter Andriof and Waddock (2002) Palazzo and Scherer (2006) Smith (2010) Hwang (2013). Kollat and Araujo (2018) state that the need for engaging stakeholders when communicating about CSR activities makes Social Network Sites like Twitter more important, given their ability to foster dialogue and content diffusion. They argue that communicating CSR on Twitter can lead to positive outcomes for companies since it is not only a way to advertise what the company does with regards to CSR, but also to receive input from other for their CSR activities. And Grunig (2009) point that social media make corporation communications more strategic, more interactive and more socially responsible. Similar work has also been pursued by Colleoni (2013) that investigates which corporate communication strategy adopted in online social media is more effective to create convergence between corporations CSR agenda and stakeholders’ social expectations, and thereby, to increase corporate legitimacy applying advanced data-mining techniques, whereas Zhang et al. (2016) explore the online debate in a CSR crisis using Twitter data relating to the recent case of the falsified Volkswagen diesel emissions that became public in 2015 focusing on capturing the issue as it evolved over time, the actors and sentiments expressed, and the responses of the organization. Findings demonstrate that CSR challenges can result in a crisis of a long duration marked by strongly expressed sentiments and a wide diversity in the views of different stakeholder groups.

Twitter can provide a favorable environment for CSR communication, and it seems that organizations with higher CSR ratings tend to build larger communities of followers on Twitter when compared to organizations with lower CSR ratings Lee et al. (2013) Kollat and Araujo (2018). But, while social media is important for CSR communication, the use of social media data is rare in the literature Chae and Park (2018) and more specifically when considering CSR in the mining sector.

Thus, this appears to serve as a reasonable framework for studying the application of sentiment analysis under CSR context where content analysis focused on this industry is scarce. These sentiments are key to go over stakeholders from different perspectives and obtain results in order to adapt CSR priorities as well as its role within mining companies, assuming that mining activities are seen as a threat to society in general.

3. Methodology

3.1. Twitter API

The Twitter Streaming API[1] allows to receive tweets and notifications in real time from Twitter. However, it requires a high performance, persistent, always in the connection between the server and Twitter.

The work of the streaming implementation is to record the incoming events as quickly as possible and process them in the background using the REST API as necessary to collect more in-depth data. The use of the REST API is subject to several speed limits by Twitter. It is important to be a responsible user of the Twitter API by planning limits to the activity within the application and monitoring the API rate limit responses. The streaming API has no speed limits since the data is sent to the server as they appear.

3.2. Streaming tweets

Fig. 1 shows the structure used for the streaming of tweets. The technologies used to this aim were the programming language Python and MongoDB.

MongoDB is a document-oriented database with dynamic schema. This allows it to offer high performance and facilitates the development of applications. In turn, it prevents from having Joins and Transactions, something very common in relational databases. It should be noted that MongoDB stores documents in JSON format and its structures.

Also, Naive Bayes sentiment classifier was applied to these tweets and the results were depicted in different graphs. Scripts were also developed to maintain the tweets collected in the database.

Tweets were streamed in real time with the API of Twitter and a string was created to filter and select the tweets according to the

[1] https://developer.twitter.com/en.html.

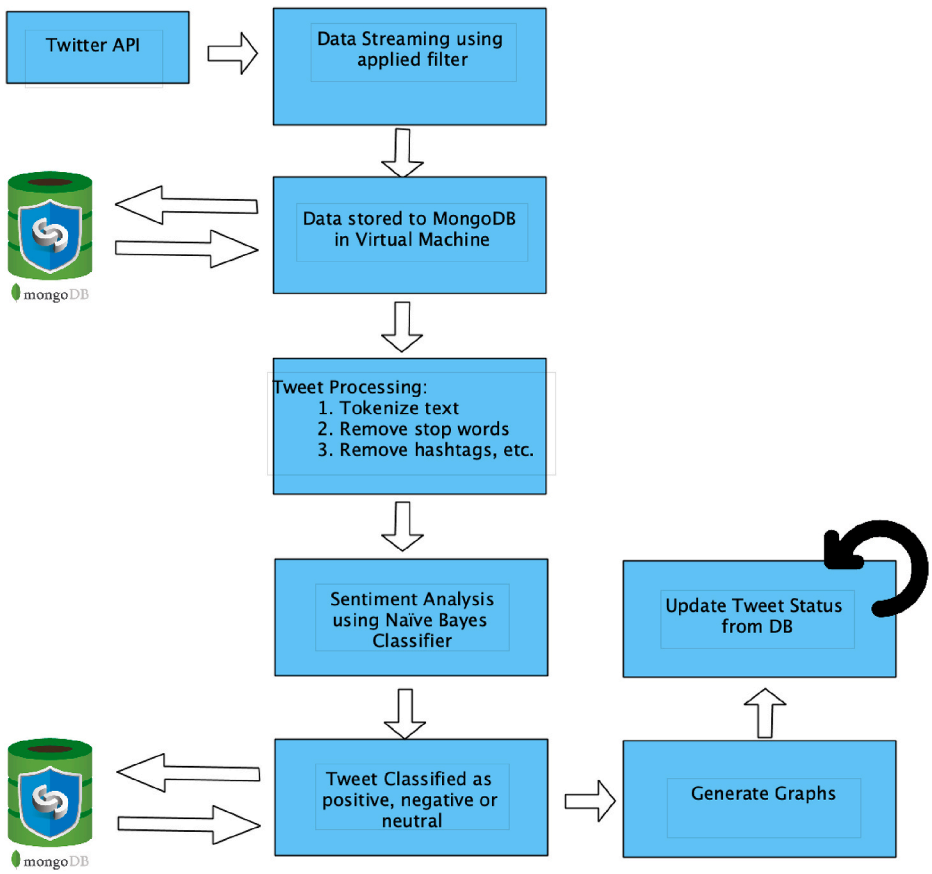

Fig. 1. Project architecture.

Table 1
Main Filters that we used to stream tweets.

Type	Filter applied
Minery Words	minery, open-pit mining, excavation, quarry, colliery
Hasthags	mining sustainability, mining purpose, responsible mining, future mining, women mining, dieselfree mining, coal mine, safetyfirst, environment, jobs, development, diversity, re-newable, safe, respect
CSR Words	CSR, ethic, ethics, ethical, responsible, responsibility, legal compliance, care, respect, reputation, social, socialgood, community, charity, philanthropy, volunteer, nonprofit, cul- ture, sustainable, sustainability, environment, renewable, climate, development, rehabilitation, stakeholder, code con- duct, eco, green, safety, safe, diversity, gender parity, trans-parency,labor practise, human rights

parameters summarized in Table 1. Each word of *Minery words* is combined with all the words of *CSR words* and with all the words of *Hashtags mining*. The result is a long list of word combinations that will ensure the tweets found are of the appropriate topic.

3.3. Algorithm

Algorithm 1 shows the system used to perform tweet streaming 24 h a day, every day of the week, to make sure we do not lose any tweets.

4. Classify algorithm

4.1. Introduction

In this section we present the method used to solve the problem of the classification of texts by their feeling at the document level. These documents are messages that have been published on the social network Twitter. A training corpus is required. Its examples were previously labeled one by one by hand with the category of feeling to which they belong (negative, positive or neutral)

Algorithm 1 Tweets Streaming Algorithm

```
procedure Stream Tweets(TwitterAPI, Filter, MongoDB)
    Initialize the connection to MongoDB
    Initialize the connection to the TwitterAPI through credentials
    Filter ← load list of words to filter while streaming
    while OnData ≠ Null do                          ▷ While we find a tweet:
        tweet = ConvertToDict(tw).lower()
        user = tweet['user']
        hashtags = tweet['hashtags']
        if tweet ∈ filter then
            if tweet ∉ MongoDB then
                MongoDB ← < tweet, user >
            end if
        else
            discard tweet
        end if
    end while
end procedure
```

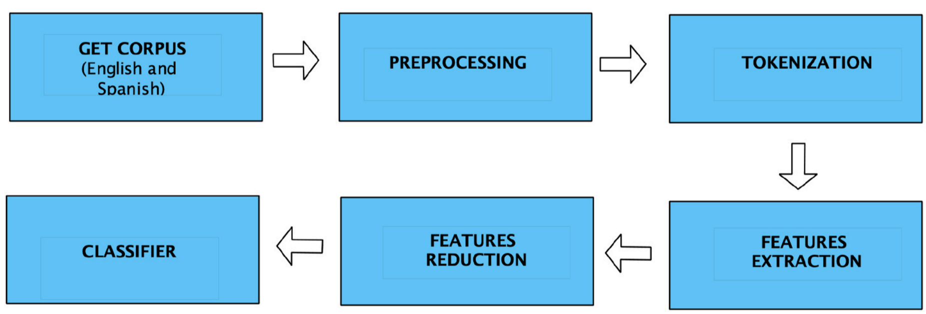

Fig. 2. Phases for training of the classifier.

4.2. Algorithm training process

As can be seen in Fig. 2, the usual construction process of a text classifier system based on machine learning consists of several sequential phases. First, it is necessary to prepare the data to train the algorithms. To do this, information must be cleaned and normalized in order to reduce or eliminate those data that may negatively influence the final result. Each of the texts undergoes then a process called tokenization, which divides them into smaller units or tokens and which are usually the words of the messages. From the tokens, the characteristics that represent the original messages are extracted and, optionally, a method can be applied to reduce their number. To finish, these characteristics are weighted according to the importance that they want to give them and with them the classifiers are trained.

Phase 1. Get corpus: In the case of tweets written in English, NLTK provides us with the corpus directly from its Python library, with more than 30,000 classified tweets, so we did not have to search for a corpus manually. In the case tweets in Spanish, we have obtained, as we mentioned previously, an XML file where we extract the manually classified tweets.

Phase 2. Preprocessing: In any method that makes use of automatic learning algorithms it is necessary to pre-treat the data with which they will be trained. The objective of this phase is to clean and normalize the information to prevent certain data from negatively influencing the final result. This question is crucial when we talk about messages extracted from social networks since it is very common to find messages with spelling mistakes, repetitions of characters, mixed uppercase and lowercase letters … For this research we have selected a set of simple rules to apply and which are common in the construction of this type of classifiers. The goal pursued is the normalization of messages, avoiding at all times that the changes applied cause the loss of the polarity of feeling.

a) Normalization of capitals and small letters: although for people it is easy to know that the words "dog" and "DOG" have exactly the same meaning, for machine learning algorithms they do not. In fact, they are treated as totally different words, without any relationship between them. To prevent this from happening, all messages are converted to their equivalent in lowercase letters.

b) Treatment of characters duplicity: in social networks it is common to repeat the same letters in words to give intensity to what is intended to express. For example, it is not the same to write “I am hungry” or “I am huuungry”. Although both sentences are conceptually equivalent, the second emphasizes the feeling of warmth by repeating the characters of the phrase. In our case, we do not need to know to what extent opinions are issued, only to know what polarity they have so it is not necessary to maintain these repetitions. Therefore, any sequence of three or more equal characters will be reduced to only two. For example, the previous case would be like “I am hungry”.
c) Elimination of line breaks: in some Twitter messages the text is written in different lines. These elements are also deleted so that the messages appear written in a single line.
d) Elimination of mentions and links: these two elements are very common in Twitter messages. The mentions are used to refer to other users of the social network by their name preceded by the symbol of the @ (for example, @RioTinto). Finally, it is possible to add to the tweets links to web pages to enrich the messages (for example: http://t.co/4diTkV2a).

Phase 3. Tokenization: Once the message normalization process is completed, the next phase is called tokenization. In this phase the texts are divided into smaller units called tokens, which normally correspond to the words of each text. This process can be as simple as separating the terms of the phrases by the white space and the punctuation characters or else consider that the grouping of certain symbols can contain some type of information that is useful to the classification process. This could be the case of emoticons, punctuation character sequences that are usually an indicator of the polarity of the feeling of the words they accompany. In our case, we will use a tokenizer that maintain the emoticons, mentions, hashtags and URLs as tokens.

Phase 4. Features Extraction: From the tokens obtained in the previous step, we will define how to represent with them the messages from which they come, thus creating the features. The usual task in text classification is to use the bag of words model (BoW) where each message is represented by its tokens without taking into account any specific order between them.

Phase 5. Features Reduction: This phase is optional and its objective is to decrease the number of characteristics of the corpus by eliminating certain tokens or converting them looking for the same way of representing them. There are two common techniques to carry out this task: stop words removal and stemming.

a) Elimination of stop words: there is a set of words that, although they are necessary to construct meaningful sentences, lack information that helps determine the polarity of the texts in which they are found. In Spanish and English, these words are prepositions, pronouns, conjunctions and the different forms of the verb “to be”, among others. By means of this technique, all the terms belonging to the stop words list will be eliminated from the model before the training of the algorithms.
b) Stemming: this is another method of morphological normalization.

In this case, a word is transformed to its root by means of the suppression of its suffixes and inflections. Following the previous example, the word “climate” would be converted to its root, “cli-mat”.

Phase 6. Classifier: Once the data is preprocessed, the Naive Classifier gets all this information and classifies the tweets according to their positivity.

4.3. Naive Bayes classifier

The Naive Bayes classifier is a probabilistic method for classification. It performs an approximate calculation of the probability that an example belongs to a class given the values of predictor variables. The simple Naive Bayes classifier is one of the most successful algorithms on many classification domains. In spite of its simplicity, it is shown to be competitive with other more complex approaches in several specific domains. This classifier learns from training data the conditional probability of each variable Xk given the class label c. Classification is then done by applying Bayes rule to compute the probability of C given the particular instance of X1,…, Xn,

$$P\ (C = c|X_1 = x1,\ \ldots,\ Xn = xn)$$

Naıve Bayes is based on the assumption that variables are conditionally independent given the class. Therefore the posterior probability of the class variable is formulated as follows,

$$P\ (C = c|X_I = x_I,\ \ldots,\ X_n = x_n)P\ (C = c)P\ (X_k = xk|C = c)$$

This equation is highly appropriate for learning from data, since the probabilities pi = P(C = ci) and pik,r = P(Xk = xrk—C = ci) may be estimated from training data. The result of the classification is the class with highest posterior probability.

In the specific case of the classification of texts, the exclusive and exhaustive events are the different classes that can be assigned to a message, so that it is not possible to assign more than one simultaneously (excluding) and those classes are all types that exist (exhaustive) Naive Bayes algorithms are often referred to as “naive” because in their calculations the characteristics selected to represent the training examples are statistically independent and contribute equally in the classification process. In other words, and in the specific case of the classification of texts, it is considered that the words of the same message do not maintain any kind of relationship with each other and the position they have within the text to which they belong is indifferent.

One of the problems of supervised methods is the need to have a representative test set (previously labeled) to train the machine learning algorithms, that is, a corpus. As reviewed above, the classification performance can directly depends on the quality of features obtained from a training dataset. Thus, our choice for Naive Bayes algorithm appears to be reasonable in terms of accuracy for our dataset characteristics according to the results obtained from Choi and Lee (2017).

In the specific case of the classification of Twitter messages, although there are currently several resources in English, it is not so easy to find them in Spanish. The creation of this type of elements is often complicated due to the enormous cost in terms of time and effort necessary to complete it. This was the problem that the authors of the first work of sentiment analysis on Twitter found Go et al. (2009), since at that time Twitter was not a popular network and, therefore, there were still no available corpora. To solve this obstacle they designed an automated system that allowed creating a corpus following the ideas presented in Read (2005). Broadly speaking, Read (2005) stated that when a user added an emoticon to a certain text, this element was an indicator of the feeling of his written words. In other words: if a text was added to a text with a happy face: (the text would have a positive connotation. If instead the symbol was a sad face:), those words would be negative.

For the tests on this research project, we used an existing corpus in Spanish whose authorship belongs to the Workshop on Semantic Analysis (TASS[2]) of the Spanish Society for the Processing of Natural

[2] http://www.sepln.org/workshops/tass/.

Language (SEPLN[3]). This society annually organizes a competition in which different methods are presented for the classification of tweets in four and six categories.

At a technical level, it is worth mentioning that the models have been written in Python[4] and that specific libraries have been used for this type of development, such as NLTK,[5] which specializes in the processing of natural languages.

The Naive Bayes algorithms are based on probabilistic theories, which allow us to estimate the probability of an event based on the probability that another event will occur, on which the first one depends. Basically, it is about estimating the probability that a document belongs to a category. Such belonging depends on the possession of a series of characteristics, of which we know the probability that they appear in the documents that belong to the category in question. Naturally, these characteristics are the terms that make up the documents and both their probability of appearance in general, and the probability that they appear in the documents of a certain category, can be obtained from the training documents. For this, the frequencies of appearance in the collection used during the learning are used.

However, there is a very important problem that obliges to expand the mentioned considerations through more sophisticated versions of mentioned algorithm. When the statistics are calculated with respect to a class, the occurrences of the words in the training documents are counted. But if a new document containing a word (term) appears that has not appeared before, the probability of the term will be zero and, consequently, also the probability of belonging to the document that contains it will be zero. with respect to any class, leading to erroneous results. The number of errors will increase when small training collections are used, therefore if we use a small number of samples, it will be more difficult a priori to consider any possible word.

Algorithm 2 gives us a brief idea of the global structure we have applied to classify tweets according to their sentiment, as well as the deletion process in case the tweet does not belong to the Spanish or English language, which are the only languages we have analyzed. It is mainly a request in our database for tweets that have not yet been classified and apply the Naive Classifier for all of them.

5. Results

Next, the main results of the research conducted are discussed.

5.1. Word clouds

After the processing of all the data collected and the stemming of the text, the roots of the words were used to get a more accurate result of the filtered words. For example, the root *"climat"* refers to all the words that are written by those letters, such as *"climate"*. These words or combination of words (i.e. *"climate"*, *"climate-change"*, *"climate-debate"*) are the ones that Twitter users use when they talk about CSR in the mining industry, which indicates that when they hear about these terms those are the first ideas with which they relate the CSR. Fig. 3 gives a general vision of which words do people use when talking about topics related to CSR in the mining industry.

Algorithm 2 Classify Tweets with Naive Bayes

```
procedure Naive Bayes(CSRp, CSRrt, SearchEngine)
    Initialize the connection to MongoDB
    NotClassified ← load all tweets without classification from MongoDB
    DataSet ← Positive Tweets + Negative Tweets
    TrainingSet ← 0.8 * Positive Tweets + 0.8 *Negative Tweets
    StopwordsEnglish ← NLTK.stopwordsenglish
    StopwordsSpanish ← TASSstopwords
    NotClassified ← Tweetsnotclassifiedf romDB
    while NotClassified /= Empty do
        if TweetLanguage = english then
            positivity ← ClassifyEnglishTweet
            MongoDB ← < positivity, tweet >
        end if
        if TweetLanguage = spanish then
            positivity ← ClassifySpanishTweet
            MongoDB ← < positivity, tweet >
        else
            delete tweet
        end if
    end while
end procedure
```

[3] http://www.sepln.org/.
[4] https://www.python.org/.
[5] http://www.nltk.org/.

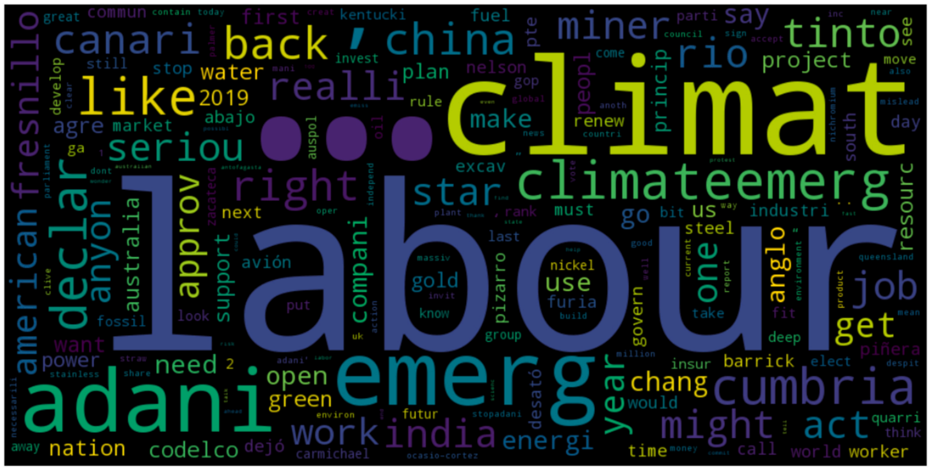

Fig. 3. Cloud of all words.

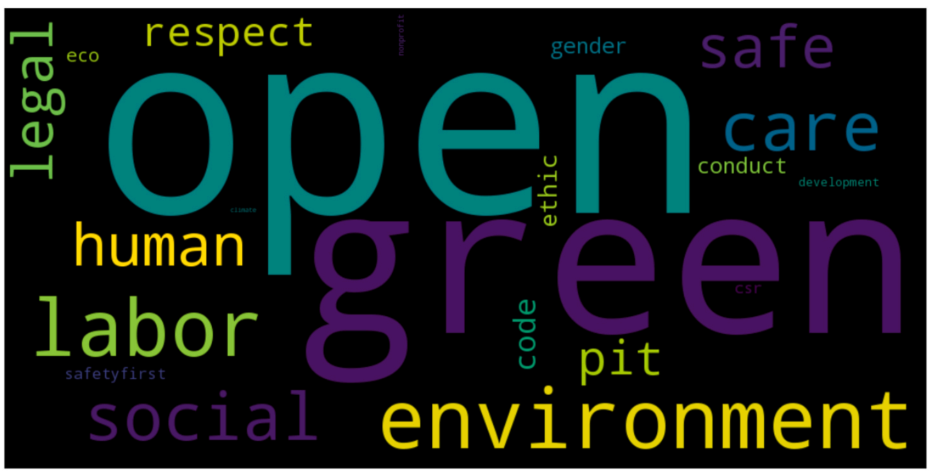

Fig. 4. Cloud of CSR words.

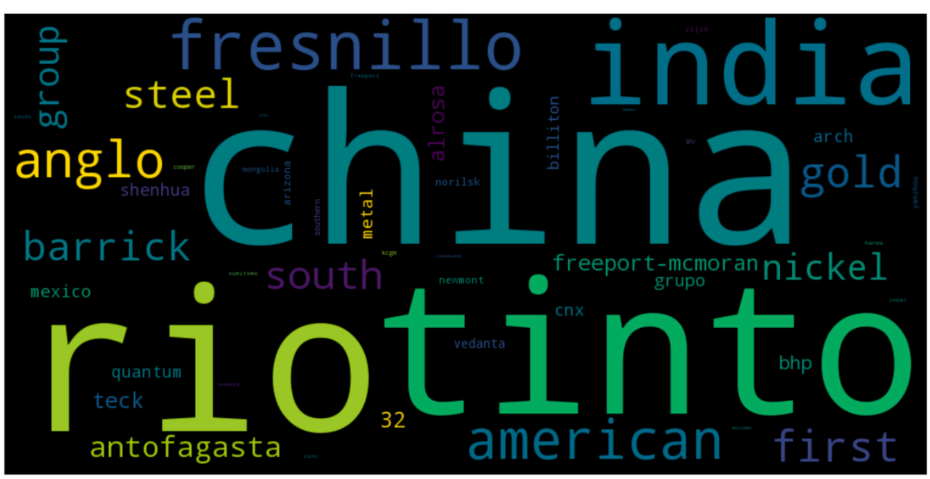

Fig. 5. Cloud of companies.

Fig. 4 gives a more detailed vision of results concerning specific CSR words. The most popular CSR terms are related to environmental issues, i.e. *"green"*, *"environment"*, *"open"* and *"eco"*, followed by words related to social issues including work or legal activity, such as *"labor"*, *"human"*, *"legal"*, *"gen-der"* and *"ethic"*. Another group of words may refer to both CSR categories: i.e. *"safe"*, *"respect"*, *"care"*, *"code"* and *"conduct"*.

These results are consistent with different studies published in the literature that reveal that one of the most visible impacts of mining is reflected in the environment Vintro et al. (2012) Jenkins and Yakovleva

(2006) and that CSR environmental actions (i.e. restoration plans, mining source reduction and control of energy consumption) are more widespread than social actions (i.e cooperation with non-governmental organizations and promotion of local communities) Vin-tro et al. (2014). This difference in the application of CSR practices as well as in the topics discussed on Twitter could be attributed to the high impact of the mining industry on its surroundings and on local communities, with morphological changes in the exploited areas, noise, dust, and surface and groundwater pollution Hilson and Murck (2000) Jenkins and Yakovleva (2006). In this sense, different tweets collected by social stakeholders (that is, tweets from anonymous personal accounts) argue that mining initiatives are still controversial when considering its impact on the land and advocate for more responsible environmental strategies:

"#stopadani India campaign must be started. Their coal mines are targetting multiple green zones. Don't destroy the land, our home for an unsustainable energy source. Ecosystem supported by the forests will never recover. water table,1000 yrs of growth #studentprotest #burning." (Tweet posted by a non- professional account)

Moreover, the top environmental words found in the Twitter analysis fit with previous results that provide a description of good environmental practices adopted by mining companies. Vintro et al. (2014) observed that investing in processes for saving energy and natural resources, reducing green-house gas emissions, improving the safety of the work environment, restoring initial environmental conditions, and minimizing the impacts to the environment were some of the most prominent examples. Suppen et al. (2006) cite different cleaner production technologies. Sandoval et al. (2006) mention transparent policies and practices to ensure long-term economic, environmental and social well-being of local communities. And Kumar (2014) reviews sustainable mining practices including environmental impact assessment, geographic information system (GIS) maps, three-dimensional (3D) models, and the use of global positions system (GPS). On this respect, different tweets posted by mining companies and collected during the analysis inform about different initiatives and projects undertaken:

"Mining companies test and treat water on at least a monthly basis and Nova Scotia Environment is the regulator that ensures companies adhere to regulations. As a result, water is usually cleaner after it has been used on a mine/quarry than it was before". (Tweet posted by a professional/official ac count)

When referring to the top social words found in the Twitter analysis, results show they are also in line with previous researches. For instance, Amponsah-Tawiah and Mensah (2016) examine the relationship and impact of occupational health and safety on employees' organizational commitment in Ghana's mining industry. Jarosławska-Sobór (2015) analyzes the social potential of a coal mining company in terms of human capital and occupational safety. Maier et al. (2014) analyze the relationship between mining and poverty, human health and the environment. And Vin-tro et al. (2012) observed that codes of conduct and professional career programs were the most popular CSR practices linked to human resources management. Results from the Twitter analysis show that ethical concerns are expected to be integrated into operations and management, which fits with similar studies (Tate et al., 2010), and issues such as human rights, child labor and working conditions are big issues that mining stakeholders care about:

"I get hurt when I see Centre and the states are ignoring the voice of tribal coal mine workers in the country". (Tweet posted by a non-professional account) "Coal miners deserve better options. Coal miners deserve better pay. Coal miners deserve safer working conditions." Tweet posted by a non-professional account)

Health-related issues are also discussed and seem to be of interest in the Twitter sphere, which is not strange considering mining has

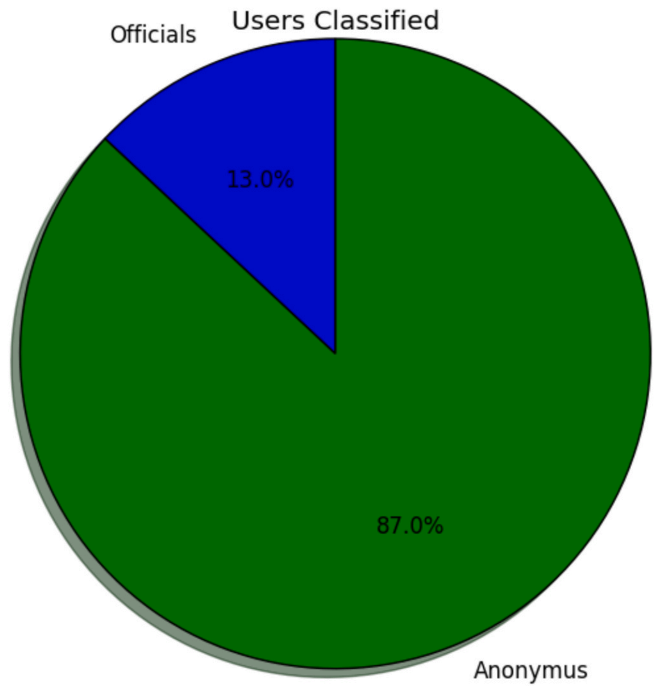

Fig. 6. Users classification by Official/Anonymous.

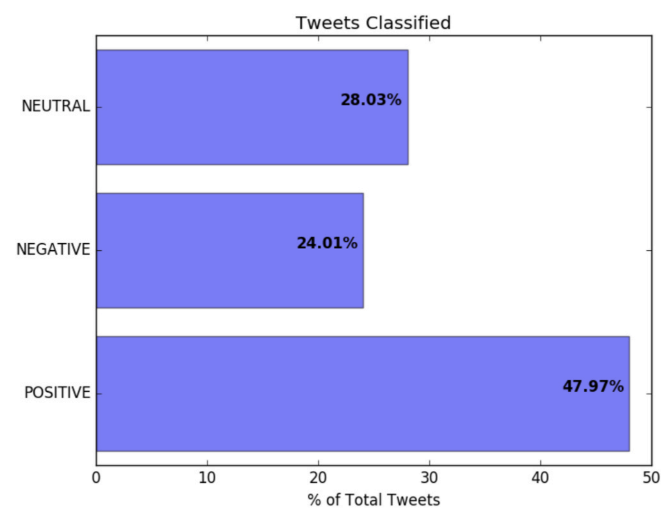

Fig. 7. Sentiment classification of the total Tweets.

traditionally been seen a dangerous activity due to its adverse social impacts when referring to health and safety problems Hilson and Murck (2000) Sánchez (1998):

"Van Howell leading the class through a workshop exercise in OSHA 7410: Managing Excavation Hazards @ Idaho Falls Safety Fest at the College of East- ern ID! The PNW OSHA Education Center was a founding partner of the Safety Fest of the Great Northwest starting in Boise in 2005! https://t.co/vf3e3VRpNa*". (Tweet posted by a professional account)*

The economic aspect of CSR is also drawn in the results. Words such as "responsibility", "plan", "invest", "profit, "company", "charity", "project", or "corporate governance" detected in the tweets, reveal that CSR is a strategic decision to create sustainable and competitive businesses. The group of words referring to the economic dimension of CSR in the mining sector turned to be less popular on the Twitter debate.

Fig. 5 shows results of words concerning profiles of companies (i.e. mineral, country, company name) with the most Twitter activity.

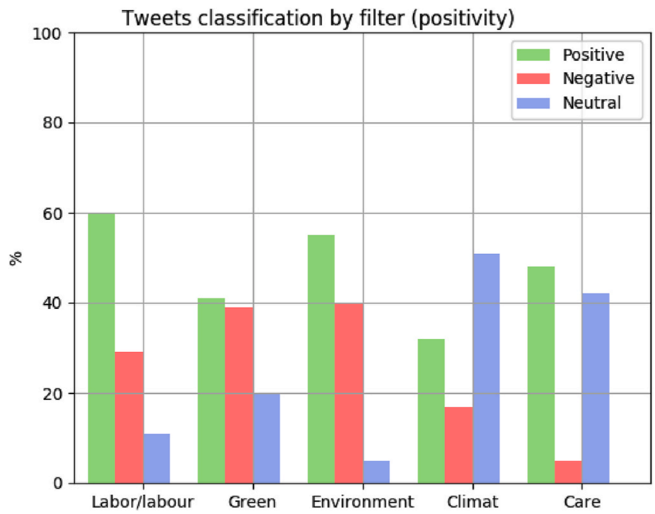

Fig. 8. Weighing with Näives Bayes of most used words.

Results reveal that companies use Twitter as a tool for promoting community-focused activities and philanthropic dimensions of CSR in order to reinforce a positive relationship with key stakeholders. CSR activities must be aligned with social norms and values to maintain and gain organizational legitimacy. Moreover, results show that interest in CSR is growing in developing countries Collier and Esteban (2007) and in countries with a bad reputation of environmental and health mining conditions Wei-ci and Chao (2011). In this line, two specific countries were popular and captured as top words: China and India. Also, different names of companies tagged are from South America. These results concur with the literature since an increasing number of general CSR studies focusing on emerging economies such as India Shirodkar et al. (2018) and specific CSR studies focusing on mining in emerging countries such as Ethiopia Woldeyohannes et al. (2018) can be found.

5.2. Users classification: official and personal anonymous twitter accounts

Fig. 6 depicts the percentage of tweets written by anonymous personal accounts (that is, public in general) and those written by official accounts (that is mining companies, mining associations and CEOs of these companies or associations). Results show that only a small amount of tweets (around 13%) are from official accounts. Official accounts generally have more followers and more influence on society than anonymous personal accounts. Therefore, mining companies and the mining sector in general should increase the adoption of new strategies to reach stakeholders satisfaction and transparent results reporting. According to a study published by Deloitte, 2016 based on tracking trends, the need for the mining sector to use social media is becoming increasingly important since it opens new opportunities for businesses to spread information and to engage stakeholders Cho et al. (2017) Kollat and Araujo (2018).

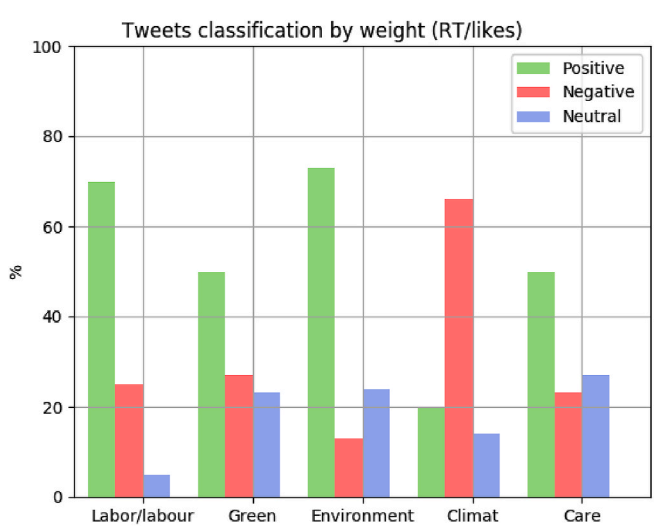

Fig. 9. Weighing with number of RT and Likes of most used words.

Table 2
Confusion matrix.

Confusion Matrix	Predicted NO	Predicted YES
Actual NO	78%	20%
Actual YES	22%	80%

Some general studies about CSR disclosure on social media can be found in the literature. For example, Chae and Park (2018) conducted a study about CSR topics and trends in Twitter and concluded that Twitter is an important channel for organizations to communicate with stakeholders. The authors found that many organizations use the social media to promote their CSR practices. Specific studies referring to the mining sector are very scarce.

5.3. Sentiment analysis of tweets

The results obtained with the user and Naive Bayes classifier (Fig. 7) show there is a slight tendency to classify the tweets as positive (around 48%). Approximately one in four tweets (25%) is considered negative, and neutral tweets account approximately for 28%.

Fig. 8 shows the results by CSR topic. We chose the 5 most popular words found in the database used for the investigation. There is not a concluding result that indicates which topics (environmental issues, social issues, …) are seen as more positive and which ones are seen as more negative.

Retweets and likes are indicators of the impact of tweets on society. In Fig. 9 we take into account retweets and likes when weighing the results. Results show more clearly which words are related to positive topics and which ones to negative. For instance, tweets with the word "environment" throw a positive sentiment while tweets with the word "climate" mainly have a negative mood. This last result fits with the study conducted by Cody et al. (2015) who found that tweets containing the word "climate" are less happy than all tweets, probably because Twitter users associate climate change with the in-crease in severity and frequency of certain environmental disasters and because the discussion on climate change on the Twitter sphere is dominated by climate change activists. On the other hand, tweets containing the word "environment" mainly refer to environmental and sustainable practices which are seen as valuable to fight against climate change and to improve the environmental behavior of companies.

5.4. Confusion matrix Naive Bayes

Table 2 shows the effectiveness of the developed Naive Bayes algorithm. Around 80% of the classifications have been made correctly, a fairly positive and effective number since most of these classifiers are around 75% effective.

6. Discussion

6.1. Implications for theory

In line with previous studies published in the literature Hilson and Murck (2000), Warhurst (2001b), Hilson and Nayee (2002), Dutta et al. (2012), we find empirical evidence that CSR in the mining sector is an increasingly studied topic by different researchers and is a core strategy of many companies.

While previous studies published in the literature have mainly used traditional methods for data collection such as reports, interviews or questionnaires, this investigation has introduced the use of big data, which opens up new opportunities for research and provides a valuable source of information. To take advantage, CSR researchers and practitioners need to be familiar with computational data collection techniques such as APIs and machine learning algorithms. Its performance strongly depends on the way of training a classification model and data properties. Twitter is the perfect data source due to maximum tweet length allowed as too long documents have noise features that hinders the sentiment analysis accuracies Choi and Lee (2017).

Our main contribution is the development of a database of tweets for the mining sector that will serve for future projects in this area, and an automatic Tweets classifier using the Naive Bayes algorithm. The results obtained here may have implications for understanding that CSR debate in the Twitter sphere is growing and that the open discussion is mainly centered on the impact of mining industries on the land and on ethical concerns referring to working conditions and occupational safety.

6.2. Implications for practice

Mining companies face a challenging competitive environment and must adapt and readapt strategies to respond positively to environmental and sustainable issues. They need to maintain a social license to operate and this is now directly linked to value perceived by stakeholders. Moreover, the digital age has turned the world and the competitive environment more transparent and companies should engage with different stakeholders in new ways. That is, the mining sector has traditionally had a business-to-business mentality, but now it must adopt a business-to-stakeholders approach.

As Lattemann and Stieglitz (2007) state, social media connects companies to their stakeholders and engage them to discuss opinions on current topics and trends. In this sense, the results of the study presented in this paper show that some mining companies have started to use social networks to provide a viable source of information and to create an environment for effective stakeholder dialogue, and different stakeholders play an active role in CSR discussions on social network sites such as Twitter.

For business, our study provides information about the main topics discussed by stakeholders which reveal the issues of concern and the perception they have about the initiatives undertaken by mining companies. This is a valuable source of information for the CSR strategy definition. The results also shed light on the need to cultivate long-term relationships with key stakeholders and to increase the use of social network sites such as Twitter for real company-stakeholder engagement and debate. It is clear that there are important benefits from knowing the main issues and concerns from stakeholders, therefore companies might redefine CSR role and priorities. This framework provides a platform for further research and application to other sectors.

For public administrations and governments, our findings seem to indicate the need for more strict environmental regulations and for actions to impulse the adoption of sustainable and ethical practices by mining companies, so that the negative impacts on the environment and society get reduced.

6.3. Limitations of the study and further research

Our research has some practical limitations, since it only considered the Tweets collected within a specific period of time and the Tweets written in English and Spanish. Moreover we developed a predefined list of specific initial parameters to filter the Tweets.

In this sense it would be interesting for further research to include more languages and more filters in order to extend the list of initial words to search within the Tweets. It would be also interesting to conduct similar studies analysing correlations between topics and words, to correlate topics and profiles of users (professional and non–professional accounts, geographical origin, …) and to use combinatorial analysis of the results with another social network to conduct a more detailed study of the CSR mining debate.

7. Conclusions

CSR communication has increased over the years, and many companies of different sectors communicate about economic, environmental and social issues in some form, and most of them engage in sustainability reporting (Frosten-son et al., 2012). Specifically, the mining industry faces pressure to adopt and communicate CSR practices.

Communicating effectively CSR activities benefits company-stakeholder engagement and develops positive attitudes from stakeholders towards the company. In this sense, some mining companies have started to use social networks to inform stakeholders and to promote their CSR practices, and now, different stakeholders play an active role in CSR discussions on social networks sites such as Twitter.

This study has aimed to understand what topics of CSR in the mining industry are discussed in the Twitter sphere. CSR appears to be an evolving construct in business and society Gupta (2015) and its dimensions and trends change over time. In this sense, the examination of big amounts of data obtained from social media may help in the aim of analyzing the CSR debate.

The results obtained in the investigation are consistent with different studies published in the literature that emphasize the open discussion about environmental issues (that is, environmental impacts of mining activities and environmental actions such as restoration plans or mining source reduction). The debate is mainly centered on the impact of mining industries on the land, and many stakeholders advocate for more environmental responsibility since they consider that mining activities are still controversial. Companies use Twitter to inform about different initiatives undertaken to be more sustainable and environmentally responsible. Social actions appear to be the second most discussed issue and the debate around it is centered in the improvement of ethical concerns considering human rights, child labor and working conditions. Occupational safety appears in many posts. The economic aspect of CSR seems to be the least commented topic in the Twitter sphere and is mainly centered on corporate governance issues.

The results show the CSR debate is increasingly growing in developing countries and in countries with a bad reputation of environmental and health mining conditions. On the other hand, the percentage of tweets posted by anonymous personal accounts (public in general) accounts for 87% while those written by official accounts (that is mining companies, mining associations and people related to these companies or associations) only reach 13%. These results reveal that the mining sector should improve the CSR disclosure and adopt a stakeholder engagement strategy grounded in the corporate stakeholder relationship perspective, the two-way symmetrical communication, and the dialogic theory of public relations Lim and Greenwood (2017). Social network sites such as Twitter can lead to positive outcomes for companies since they are not only a way to advertise what the company does with regards to CSR, but also to receive input from other for their CSR activities. Different authors defend that social media make corporation communications more strategic, more interactive and more socially responsible Grunig (2009).

Results also reveal a slight tendency to positive tweets (48%) in front of negative tweets (25%) and neutral tweets (28%). There is not a concluding result that indicates whether environmental issues, social issues, or economic issues are seen as more positive and which ones are seen as more negative. For instance, when considering environmental issues, tweets with the word "environment" have a positive mood and are aligned with the value of environmental and sustainable practices undertaken by mining companies. On the other hand, tweets with the word "climate" are less happy and associate climate change with the increase in severity and frequency of certain environmental disasters.

Author statement

Adrià Pons (Ph.D. Student in Business Administration)

Adrià is a PhD student and this work is part of his thesis. He participated in the analysis of data and in writing the manuscript. He also revised and edited the document.

Carla Vintrò (Ph.D. in Natural Resources and Environment)

Carla was responsible of the conceptualization and formal analysis. She supervised the analysis and interpretation of data and participated in writing the manuscript.

Josep Rius (Ph.D. in Computer Science)

Josep designed the methodology and the software used to connect with twitter and collect the data. Together with Carla, he also supervised Adrià in all phases and participate in writing the manuscript.

Jordi Vilaplana (Ph.D. in Computer Science)

Jordi was the main developer of the software tool used to collect the analyzed data. Also, helped creating scripts to manage and get the results from this data. He also participated in writing the manuscript.

Acknowledgments

This work was supported by the Ministerio de Economía y Competitividad under contract TIN2017-84553-C2-2-R. Also, the authors are members of the research group 2017-SGR363, funded by the Generalitat de Catalunya. Finally, we thank Victor Ferrer for his contribution.

References

Amponsah-Tawiah, K., Mensah, J., 2016. Occupational health and safety and organizational commitment: evidence from the ghanaian mining industry. Saf. Health Work 7, 225–230. https://doi.org/10.1016/j.shaw.2016.01.002.

Andriof, J., Waddock, S., 2002. Unfolding stakeholder engagement. In: Unfolding Stakeholder Thinking: Theory, Responsibility and Engagement, pp. 19–42.

Barnes, N., M, E., 2010. The fortune 500 and social media: a longitudinal study of blogging and twitter usage by America's largest companies. In: UMass Dartmouth Center for Marketing Research, vol. 6, p. 2011.

boyd, d., Golder, S., Lotan, G., 2010. Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter, pp. 1–10. https://doi.org/10.1109/HICSS.2010.412.

Carroll, A., 1991. The pyramid of corporate social responsibility: toward the moral management of organizational stakeholders. Bus. Horiz. 34, 39–48. https://doi.org/10.1016/0007-6813(91)90005-G.

Chae, B., Park, E., 2018. Corporate social responsibility (csr): a survey of topics and trends using twitter data and topic modeling. Sustainability 10, 2231. https://doi.org/10.3390/su10072231.

Cho, M., Furey, L.D., Mohr, T., 2017. Communicating corporate social respon- sibility on social media: strategies, stakeholders, and public engagement on corporate facebook. Bus. Prof. Commun. Q. 80, 52–69. https://doi.org/10.1177/2329490616663708.

Choi, Y., Lee, H., 2017. Data properties and the performance of sentiment clas- sification for electronic commerce applications. Inf. Syst. Front 19, 993–1012.

Claver Cortés, E., Molina-Azorin, J., Dolores López Gomer, M., Zaragoza-Sáez, P., 2004. La integraci'on del capital medioambiental en el capital intelectual de la empresa. Revista de economía y empresa, pp. 11–28. ISSN 0213-2834 21.

Cody, E.M., Reagan, A.J., Mitchell, L., Dodds, P.S., Danforth, C.M., 2015. Climate change sentiment on twitter: an unsolicited public opinion poll. PloS One 10, 1–18. https://doi.org/10.1371/journal.pone.0136092.

Colleoni, E., 2013. CSR communication strategies for organizational legitimacy in social media. Corp. Commun. Int. J. 18, 228–248. https://doi.org/10.1108/13563281311319508.

Collier, J., Esteban, R., 2007. Corporate social responsibility and employee commitment. Bus. Ethics Eur. Rev. 16, 19–33. https://doi.org/10.1111/j.1467-8608.2006.00466.x.

Cortes, C., Jackel, L.D., Chiang, W.P., 1995. Limits on learning machine accu- racy imposed by data quality. In: Advances in Neural Information Processing Systems, pp. 239–246.

Crane, L., Self, R., 2014. Big data analytics: a threat or an opportunity for knowledge management?. In: Lecture Notes in Business Information Processing 185, pp. 25–34. https://doi.org/10.1007/978-3-319-08618-7_3.

Dahlsrud, A., 2008. How corporate social responsibility is defined: an anal- ysis of 37 definitions. Corp. Soc. Responsib. Environ. Manag. 15, 1–13. https://doi.org/10.1002/csr.132.

Dellarocas, C., 2003. The digitization of word of mouth: promise and challenges of online feedback mechanisms. Manag. Sci. 49, 1275–1444. https://doi.org/10.2139/ssrn.393042.

Deloitte, 2016. The top 10 issues mining companies will face in the coming year. Track. Trends 50.

Dutta, S., Lawson, R., Marcinko, D., 2012. Paradigms for sustainable develop- ment: implications of management theory. Corp. Soc. Responsib. Environ. Manag. 19, 1–10. https://doi.org/10.1002/csr.259.

Escanciano, C., Fernández Muñiz, B., Sánchez, A., 2010. Organizaci'on de la actividad preventiva y gesti'on de la seguridad y salud laboral en la miner'ıa española: experiencia de las empresas certificadas iso 9001. Dir. Organ. 40, 86–98.

Fonseca, A., 2010. How credible are mining corporations' sustainability reports? a critical analysis of external assurance under the requirements of the inter- national council on mining and metals. Corp. Soc. Responsib. Environ. Manag. 17, 355–370. https://doi.org/10.1002/csr.230.

Go, A., Bhayani, R., Huang, L., 2009. Twitter Sentiment Classification Using Distant Supervision. CS224N Project Report, Stanford 1, 2009.

Grunig, J., 2009. Paradigms of global public relations in an age of digitalisation. PRism 6, 1–19.

Gupta, S., 2015. CSR - Summarizing Evolution of the Concept. 4th Annual International Commerce Convention, vol. 4. Dept. of Commerce, Delhi School of Economics, University of Delhi, India, p. 18. https://doi.org/10.13140/RG.2.1.1053.4004.

Hamann, R., 2004. Corporate social responsibility, partnerships, and institu- tional change: the case of mining companies in South Africa. Nat. Resour. Forum 28, 278–290. https://doi.org/10.1111/j.1477-8947.2004.00101.x.

Hassan, A., Ibrahim, E., 2012. Corporate environmental information disclosure: factors influencing companies' success in attaining environmental awards. Corp. Soc. Responsib. Environ. Manag. 19, 32–46. https://doi.org/10.1002/csr.278.

Hilson, G., Murck, B., 2000. Sustainable development in the mining industry: clarifying the corporate perspective. Resour. Pol. 26, 227–238. https://doi.org/10.1016/S0301-4207(00)00041-6.

Hilson, G., Nayee, V., 2002. Environmental management system implementa- tion in the mining industry: a key to achieving cleaner production. Int. J. Miner. Process. 64, 19–41. https://doi.org/10.1016/S0301-7516(01)00071-0.

Hwang, S., 2013. The effect of twitter use on politicians' credibility and attitudes toward politicians. J. Publ. Relat. Res. 25, 246–258. https://doi.org/10.1080/1062726X.2013.788445.

Jarosławska-Sobór, S., 2015. Social potential growth of a mining company on the basis of human capital and occupational safety. J. Sustain. Min. 14, 195–202. https://doi.org/10.1016/j.jsm.2016.02.002.

Jenkins, H., Yakovleva, N., 2006. Corporate social responsibility in the mining industry: exploring trends in social and environmental disclosure. J. Clean. Prod. 14, 271–284. https://doi.org/10.1016/j.jclepro.2004.10.004.

Kelleher, T., 2007. Public Relations Online: Lasting Concepts for Changing Media. SAGE Publications, Inc. https://doi.org/10.4135/9781452225876.

Kira, K., Rendell, L.A., 1992. A practical approach to feature selection. In: Machine Learning Proceedings 1992. Elsevier, pp. 249–256.

Kollat, J., Araujo, T., 2018. Communicating effectively about csr on twitter: the power of engaging strategies and storytelling elements. Internet Res. 28, 419–431. https://doi.org/10.1108/IntR-04-2017-0172.

Kühl, N., Goutier, M., Ensslen, A., Jochem, P., 2019. Literature vs. twitter: em- pirical insights on customer needs in e-mobility. J. Clean. Prod. 213, 508–520. https://doi.org/10.1016/j.jclepro.2018.12.003.

Kumar, N.P., 2014. Review on sustainable mining practices. Int. Res. J. Earth Sci. 2, 26–29.

Lattemann, C., Stieglitz, S., 2007. Online Communities for Customer Relationship Management on Financial Stock Markets - a Case Study from a Project at the Berlin Stock Exchange, p. 76.

Lee, S.Y., Carroll, C., 2011. The emergence, variation, and evolution of corporate social responsibility in the public sphere, 1980–2004: the expo- sure of firms to public debate. J. Bus. Ethics 104, 115–131. https://doi.org/10.1007/s10551-011-0893-y.

Lee, K., Oh, W.Y., Kim, N., 2013. Social media for socially responsible firms: analysis of fortune 500's twitter profiles and their csr/csir ratings. J. Bus. Ethics 118, 791–806. https://doi.org/10.1007/s10551-013-1961-2.

Lim, J.S., Greenwood, C.A., 2017. Communicating corporate social responsi- bility (csr): stakeholder responsiveness and engagement strategy to achieve csr goals. Publ. Relat. Rev. 43, 768–776. https://doi.org/10.1016/j.pubrev.2017.06.007.

Maier, R., Díaz-Barriga, F., Field, J., Hopkins, J., Klein, B., Poulton, M., 2014. Socially responsible mining: the relationship between mining and poverty, human health and the environment. Rev. Environ. Health 29, 791–806. https://doi.org/10.1515/reveh-2014-0022.

Marwick, A., Boyd, D., 2011. To see and be seen: celebrity practice on twit- ter. Convergence. Int. J. Res. New Media Technol. 17, 139–158. https://doi.org/10.1177/1354856510394539.

Mclaren, S., Wehrmeyer, W., Argust, P.W., Graham, S., Robertson, J., 1999. Sustainability and the primary extraction industries: theories and practice. Resour. Pol. 25, 277–286. https://doi.org/10.1016/S0301-4207(00)00003-9.

Mendes, L., Rodrigues, M., 2018. Mapping of the literature on social respon- sibility in the mining industry: a systematic literature review. J. Clean. Prod. 181, 88–101. https://doi.org/10.1016/j.jclepro.2018.01.163.

Mitchell, T.M., 1999. Machine learning and data mining. Commun. ACM 42, 30–36.

Mudd, G.M., 2010. The environmental sustainability of mining in Australia: key mega-trends and looming constraints. Resour. Pol. 35, 98–115. https://doi.org/10.1016/j.resourpol.2009.12.001.

Newbold, J., 2006. Chile's environmental momentum: iso 14001 and the large- scale mining industry – case studies from the state and private sector. J. Clean. Prod. 14, 248–261. https://doi.org/10.1016/j.jclepro.2004.05.010.

Palazzo, G., Scherer, A., 2006. Corporate legitimacy as deliberation: a com- municative framework. J. Bus. Ethics 66, 71–88. https://doi.org/10.1007/s10551-006-9044-2.

Raufflet, E., Barin Cruz, L., Brès, L., 2014. An assessment of corporate social responsibility practices in the mining and oil and gas industries. J. Clean. Prod. 84, 256–270. https://doi.org/10.1016/j.jclepro.2014.01.077.
Read, J., 2005. Using emoticons to reduce dependency in machine learning tech- niques for sentiment classification. Assoc. Comput. Ling. 43–48.
Rodrigues da Silva Enríquez, M.A., Drummond, J., 2007. Social-environmental certification: sustainable development and competitiveness in the mineral industry of the brazilian amazon. Nat. Resour. Forum 31, 71–86. https://doi.org/10.1111/j.1477-8947.2007.00127.x.
Rybalko, S., Seltzer, T., 2010. Dialogic communication in 140 characters or less: how fortune 500 companies engage stakeholders using twitter. Publ. Relat. Rev. 36, 336–341. https://doi.org/10.1016/j.pubrev.2010.08.004.
Sánchez, L., 1998. Industry response to the challenge of sustainability: the case of the canadian nonferrous mining sector. Environ. Manag. 22, 521–531.
Sandoval, M.C., Veiga, M.M., Hinton, J., Sandner, S., 2006. Application of sustainable development concepts to an alluvial mineral extraction project in lower caroni river, Venezuela. J. Clean. Prod. 14, 415–426. https://doi.org/10.1016/j.jclepro.2004.10.007. improving Environmental, Economic and Ethical Performance in the Mining Industry. Part 1. Environmental Management and Sustainable Development.
Sanmiquel, L., Rossell Garriga, J., Vintro, C., 2015. Study of Spanish mining accidents using data mining techniques. Saf. Sci. 75, 49–55. https://doi.org/10.1016/j.ssci.2015.01.016.
Shirodkar, V., Beddewela, E., Richter, U.H., 2018. Firm-level determinants of political csr in emerging economies: evidence from India. J. Bus. Ethics 148, 673–688. https://doi.org/10.1007/s10551-016-3022-0.
Sinding, K., 2009. Environmental impact assessment and management in the mining industry. Nat. Resour. Forum 23, 57–63. https://doi.org/10.1111/j.1477-8947.1999.tb00238.x.
Smith, B.G., 2010. Socially distributing public relations: twitter, Haiti, and interactivity in social media. Publ. Relat. Rev. 36, 329–335. https://doi.org/10.1016/j.pubrev.2010.08.005.
Stelzner, M., 2012. 2012 Social Media Marketing Industry Report, How Mar- Keters Are Using Social Media to Grow Their Businesses. Social Media Examiner. Technical Report.
Suppen, N., Carranza, M., Huerta, M., Hernández, M.A., 2006. Environ- mental management and life cycle approaches in the mexican mining in- dustry. J. Clean. Prod. 14, 1101–1115. https://doi.org/10.1016/j.jclepro.2004.12.020.
Tate, W.L., Ellram, L.M., Kirchoff, J.F., 2010. Corporate social re-sponsibility reports: a thematic analysis related to supply chain management. J. Supply Chain Manag. 46, 19–44. https://doi.org/10.1111/j.1745-493X.2009.03184.x.
Tredinnick, L., 2006. Web 2.0 and business: a pointer to the intranets of the future? Bus. Inf. Rev. 23, 228–234. https://doi.org/10.1177/0266382106072239.
Türkel, S., Akan, A., 2015. Corporate Social Responsibility (CSR) Communi- Cation. A Turkish Industry Example, London, pp. 151–174. https://doi.org/10.1057/9781137388551_7.
Van Zyl, D., Lohry, J., Reid, R., 2007. Evaluation of resource management plans in Nevada using seven questions to sustainability. In: Proceedings of the 3rd International Conference on Sustainable Development Indicators in the Mineral Industries, pp. 403–410.
Vintro, C., Fortuny-Santos, J., Sanmiquel, L., Freijo, M., Edo, J., 2012. Is corpo- rate social responsibility possible in the mining sector? Evidence from Catalan companies. Resour. Pol. 37, 118–125. https://doi.org/10.1016/j.resourpol.2011.10.003.
Vintro, C., Sanmiquel, L., Freijo, M., 2014. Environmental sustainability in the mining sector:evidence from Catalan companies. J. Clean. Prod. 84, 155–163. https://doi.org/10.1016/j.jclepro.2013.12.069.
Warhurst, A., 2001a. Corporate citizenship and corporate social investment drivers of trisector partnerships. J. Comput. Chem. 1, 57–73.
Warhurst, A., 2001b. Corporate citizenship and corporate social investment drivers of trisector partnerships. J. Comput. Chem. 1, 57–73.
Warhurst, A., Noronha, L., 2009. Corporate strategy and viable future land use: planning for closure from the outset of mining. Nat. Resour. Forum 24, 153–164. https://doi.org/10.1111/j.1477-8947.2000.tb00939.x.
Waters, R., Jamal, J., 2011. Tweet, tweet, tweet: a content analysis of nonprofit organizations' twitter updates. Publ. Relat. Rev. 37, 321–324. https://doi.org/10.1016/j.pubrev.2011.03.002.
Wei-ci, G., Chao, W., 2011. Comparative study on coal mine safety between China and the us from a safety sociology perspective. Procedia Eng. 26, 2003–2011. https://doi.org/10.1016/j.proeng.2011.11.2397 iSMSSE2011.
Wheeler, D., Fabig, H., Boele, R., 2002. Paradoxes and dilemmas for stake- holder responsive firms in the extractive sector: lessons from the case of shell and the ogoni. J. Bus. Ethics 39, 297–318. https://doi.org/10.1023/A:1016542207069.
Woldeyohannes, D.Y., Salem, H., Ayongaba, B., Veljkovic, Z., 2018. Mining sector challenges in developing countries, tigray, Ethiopia and inspirational success stories from Australia. Int. J. Min. Miner. Eng. 9, 321–367. https://doi.org/10.1504/IJMME.2018.10018510.
Xiang, Z., Gretzel, U., 2010. Role of social media in online travel information search. Tourism Manag. 31, 179–188. https://doi.org/10.1016/j.tourman.2009.02.016.
Zhang, B., Marita, V., Veijalainen, J., Wang, S., Kotkov, D., 2016. The issue arena of a corporate social responsibility crisis-the volkswagen case in twitter. Stud. Media Commun. 4, 32–43.
Zhu, Q., Geng, Y., Fujita, T., Hashimoto, S., 2010. Green supply chain management in leading manufacturers: case studies in Japanese large companies. Manag. Res. Rev. 33, 380–392. https://doi.org/10.1108/01409171011030471.